

#5

OIPE

RAW SEQUENCE LISTING

DATE: 10/15/2001

PATENT APPLICATION: US/09/900,038A

TIME: 13:34:08

Input Set : A:\766.53 Sequence Listing.txt

Output Set: N:\CRF3\10152001\I900038A.raw

3 <110> APPLICANT: Miyake, Katsuhide
4 Watanabe, Masaki
5 Iijima, Shinji
7 <120> TITLE OF INVENTION: Beta 1,3-galactosyltransferase and DNA encoding the same
9 <130> FILE REFERENCE: 766.53
11 <140> CURRENT APPLICATION NUMBER: US 09/900,038A
C--> 12 <141> CURRENT FILING DATE: 2001-09-21
14 <150> PRIOR APPLICATION NUMBER: JP 2001-392
15 <151> PRIOR FILING DATE: 2001-01-05
17 <160> NUMBER OF SEQ ID NOS: 8
19 <170> SOFTWARE: PatentIn Ver. 2.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 313
23 <212> TYPE: PRT
24 <213> ORGANISM: Streptococcus agalactiae Type Ib
W--> 25 <400> SEQUENCE: 1
26 Met Asn Tyr Ser Ile Ile Met Ser Val Tyr Asn Glu Pro Leu Asn Tyr
27 1 5 10 15
29 Val Arg Asp Ser Val Glu Ser Ile Leu Asn Gln Thr Leu Thr Asp Phe
30 20 25 30
32 Glu Phe Ile Ile Val Ile Asp Asn Pro Ser Arg Gly Asp Leu Lys Gln
33 35 40 45
35 Phe Leu Thr Glu Tyr Ser Val Val Asp Asn Arg Ile Lys Ile Leu Leu
36 50 55 60
38 Asn Glu Glu Asn Ile Gly Leu Ala Ser Ser Leu Asn Lys Ala Val Lys
39 65 70 75 80
41 Ile Ser Lys Gly Glu Tyr Ile Phe Arg Met Asp Ala Asp Asp Ile Ser
42 85 90 95
44 Tyr Pro Ser Arg Phe Asp Lys Gln Ile Arg Phe Met Glu Glu Asn Ser
45 100 105 110
47 Leu Asp Phe Ser Ala Thr Leu Ile Glu Leu Ile Asp Gln Lys Gly Asn
48 115 120 125
50 Leu Val Tyr Lys Lys Gln Arg Glu Ser Asn Lys Ile Tyr Leu Thr Asn Asp
51 130 135 140
53 Ile Arg Lys Met Leu Leu Asn Arg Ser Ile Leu Ala His Pro Thr Trp
54 145 150 155 160
56 Cys Val Lys Lys Lys Val Phe Asp Lys Leu Met Gly Tyr Arg Asp Leu
57 165 170 175
59 Val Pro Val Glu Asp Tyr Asp Phe Ala Ile Arg Gly Ala Leu Ala Asp
60 180 185 190
62 Phe Lys Ile Gly Leu Leu Asn Lys Val Leu Leu Gln Tyr Arg Leu Asn
63 195 200 205
65 Glu Asn Gly Ile Ser Gln Thr Asn Lys Phe Lys Gln Tyr Ile Tyr Ser
66 210 215 220
68 Ala Ile Leu Gln Asp Phe Tyr Lys Glu Lys Ser Tyr Ile Asp Ile Thr
69 225 230 235 240
71 Lys Ile Thr Asn Tyr Phe Gln Glu Tyr Val Ile Lys Lys Arg Tyr Thr

ENTERED

RAW SEQUENCE LISTING

DATE: 10/15/2001

PATENT APPLICATION: US/09/900,038A

TIME: 13:34:08

Input Set : A:\766.53 Sequence Listing.txt

Output Set: N:\CRF3\10152001\I900038A.raw

```

72          245          250          255
74 Gln Gln Glu Leu Ser Lys Tyr Phe Glu Leu Lys Ser Thr Pro Ser Ile
75          260          265          270
77 Thr Ile Arg Lys Leu Tyr Ile Cys Leu Tyr Leu Tyr Phe Lys Ser Pro
78          275          280          285
80 Leu Val Arg Arg Leu Leu Ile Asn Asp Ile Asn Ile Leu Val Leu Lys
81          290          295          300
83 Leu Phe Gly Gly Glu Lys Gln Ser Asp
84 305          310
86 <210> SEQ ID NO: 2
87 <211> LENGTH: 939
88 <212> TYPE: DNA
89 <213> ORGANISM: Streptococcus agalactiae Type Ib
W--> 90 <400> SEQUENCE: 2
91 atg aat tat agt atc att atg tcg gta tat aat gag cct tta aat tat 48
92 Met Asn Tyr Ser Ile Ile Met Ser Val Tyr Asn Glu Pro Leu Asn Tyr
93 1 5 10 15
95 gtg aga gat tca gta gaa tct ata tta aat caa acg ctt act gat ttt 96
96 Val Arg Asp Ser Val Glu Ser Ile Leu Asn Gln Thr Leu Thr Asp Phe
97 20 25 30
99 gag ttc ata att gtc att gat aat cca agt aga ggt gat tta aag caa 144
100 Glu Phe Ile Ile Val Ile Asp Asn Pro Ser Arg Gly Asp Leu Lys Gln
101 35 40 45
103 ttc tta aca gaa tat tca gtt gta gat aat aga ata aaa atc ttg ctt 192
104 Phe Leu Thr Glu Tyr Ser Val Val Asp Asn Arg Ile Lys Ile Leu Leu
105 50 55 60
107 aat gaa gaa aat att ggt tta gca tca agt ttg aac aaa gcg gtg aaa 240
108 Asn Glu Glu Asn Ile Gly Leu Ala Ser Ser Leu Asn Lys Ala Val Lys
109 65 70 75 80
111 att tct aag gga gaa tat att ttt aga atg gat gct gat gat att tca 288
112 Ile Ser Lys Gly Glu Tyr Ile Phe Arg Met Asp Ala Asp Asp Ile Ser
113 85 90 95
115 tat cca agt aga ttt gat aag caa att cgt ttt atg gag gaa aat tca 336
116 Tyr Pro Ser Arg Phe Asp Lys Gln Ile Arg Phe Met Glu Glu Asn Ser
117 100 105 110
119 ttg gat ttc tca gca act cta ata gaa ttg ata gac caa aaa gga aat 384
120 Leu Asp Phe Ser Ala Thr Leu Ile Glu Leu Ile Asp Gln Lys Gly Asn
121 115 120 125
123 tta gta tat aaa caa cga gaa agt aat aaa ata tac tta act aat gat 432
124 Leu Val Tyr Lys Gln Arg Glu Ser Asn Lys Ile Tyr Leu Thr Asn Asp
125 130 135 140
127 ata cgg aag atg tta ttg aat aga tct ata ctt gcc cac cca acg tgg 480
128 Ile Arg Lys Met Leu Leu Asn Arg Ser Ile Leu Ala His Pro Thr Trp
129 145 150 155 160
131 tgc gta aaa aag aaa gtt ttc gat aag tta atg gga tat aga gat tta 528
132 Cys Val Lys Lys Lys Val Phe Asp Lys Leu Met Gly Tyr Arg Asp Leu
133 165 170 175
135 gta cct gtt gaa gat tat gat ttt gca ata aga gga gct ctg gct gat 576
136 Val Pro Val Glu Asp Tyr Asp Phe Ala Ile Arg Gly Ala Leu Ala Asp

```

RAW SEQUENCE LISTING

DATE: 10/15/2001

PATENT APPLICATION: US/09/900,038A

TIME: 13:34:08

Input Set : A:\766.53 Sequence Listing.txt

Output Set: N:\CRF3\10152001\I900038A.raw

```

137          180          185          190
139 ttc aaa atc ggc tta ctc aat aaa gta ctt tta cag tat aga tta aac 624
140 Phe Lys Ile Gly Leu Leu Asn Lys Val Leu Leu Gln Tyr Arg Leu Asn
141          195          200          205
143 gag aat gga ata tca caa acc aat aag ttt aag caa tat att tac tca 672
144 Glu Asn Gly Ile Ser Gln Thr Asn Lys Phe Lys Gln Tyr Ile Tyr Ser
145          210          215          220
147 gct att tta caa gat ttt tat aaa gaa aaa tct tat att gat atc aca 720
148 Ala Ile Leu Gln Asp Phe Tyr Lys Glu Lys Ser Tyr Ile Asp Ile Thr
149 225          230          235          240
151 aaa att act aat tac ttt caa gag tat gtg ata aag aaa cgc tat act 768
152 Lys Ile Thr Asn Tyr Phe Gln Glu Tyr Val Ile Lys Lys Arg Tyr Thr
153          245          250          255
155 cag caa gag ctc tct aaa tat ttt gag cta aaa tct acc cct agt att 816
156 Gln Gln Glu Leu Ser Lys Tyr Phe Glu Leu Lys Ser Thr Pro Ser Ile
157          260          265          270
159 act att aga aaa cta tat att tgt tta tat tta tac ttt aag tct ccc 864
160 Thr Ile Arg Lys Leu Tyr Ile Cys Leu Tyr Leu Tyr Phe Lys Ser Pro
161          275          280          285
163 ttg gtt agg agg tta tta ata aat gat att aat att tta gta ctg aaa 912
164 Leu Val Arg Arg Leu Leu Ile Asn Asp Ile Asn Ile Leu Val Leu Lys
165          290          295          300
167 ttg ttt gga gga gag aaa caa agt gac 939
168 Leu Phe Gly Gly Glu Lys Gln Ser Asp
169 305          310
172 <210> SEQ ID NO: 3
173 <211> LENGTH: 6865
174 <212> TYPE: DNA
175 <213> ORGANISM: Streptococcus agalactiae type Ib
177 <220> FEATURE:
178 <221> NAME/KEY: CDS
179 <222> LOCATION: (617)..(1789)
181 <220> FEATURE:
182 <221> NAME/KEY: CDS
183 <222> LOCATION: (1816)..(2262)
185 <220> FEATURE:
186 <221> NAME/KEY: CDS
187 <222> LOCATION: (2265)..(2744)
189 <220> FEATURE:
190 <221> NAME/KEY: CDS
191 <222> LOCATION: (2843)..(3979)
193 <220> FEATURE:
194 <221> NAME/KEY: CDS
195 <222> LOCATION: (3982)..(4953)
197 <220> FEATURE:
198 <221> NAME/KEY: CDS
199 <222> LOCATION: (5009)..(5947)
201 <400> SEQUENCE: 3
202 agatcttggga gatattatct gtgaaaccaa tgttcctaga ctgatggtcg ttccttcagg 60

```

RAW SEQUENCE LISTING

DATE: 10/15/2001

PATENT APPLICATION: US/09/900,038A

TIME: 13:34:08

Input Set : A:\766.53 Sequence Listing.txt

Output Set: N:\CRF3\10152001\I900038A.raw

```

204 gaaagtacca ccaaattccaa cagcattact tcagaacgct tattttaata agatgattga 120
206 agctattataa aatatatttg attatattat catcgatact ccacctattg gtttagttgt 180
208 tgatgccgca ataatcgcta atgcttgcga tggttttatt ttagtaaccc aagcaggtag 240
210 aataaaacgt aattatgttg aaaaagcaaa agaacagatg gaacaaagtg gttcaaagtt 300
212 cttagggtatt attcttaata aagttaatga atctgttgct acttacggcg attatggaaa 360
214 ttacggaaaa agggatagaa aaaggaagta aggggctctt gtattgaaag aaaaagaaaa 420
216 tatacaaaaag attattatag cgatgattca aaccgttggt gtttattttt ctgcaagttt 480
218 gacattaaca ttaattactc ccaactttaa aagcaataaa gatttattgt ttgttctatt 540
220 gatacattat attgtctttt atctttctga tttttacaga gacttttgga gtcgtggcta 600
222 tcttgaagag tttaaa atg gta ttg aaa tac agc ttt tac tat att ttc ata 652
223 Met Val Leu Lys Tyr Ser Phe Tyr Tyr Ile Phe Ile
224 1 5 10
226 tca agt tca tta ttt ttt att tct aaa aac tct ttt aca acg aca cga 700
227 Ser Ser Ser Leu Phe Phe Ile Ser Lys Asn Ser Phe Thr Thr Thr Arg
228 15 20 25
230 ctt tcc ttt ttt act ttt att gct atg aat tcg att tta tta tat cta 748
231 Leu Ser Phe Phe Thr Phe Ile Ala Met Asn Ser Ile Leu Leu Tyr Leu
232 30 35 40
234 ttg aat tca ttt tta aaa tat tat cga aaa tat tct tac gct aag ttt 796
235 Leu Asn Ser Phe Leu Lys Tyr Tyr Arg Lys Tyr Ser Tyr Ala Lys Phe
236 45 50 55 60
238 tca cga gat acc aaa gtt gtt ttg ata acg aat aag gat tct tta tca 844
239 Ser Arg Asp Thr Lys Val Val Leu Ile Thr Asn Lys Asp Ser Leu Ser
240 65 70 75
242 aaa atg acc ttt agg aat aaa tac gac cat aat tat atc gct gtc tgt 892
243 Lys Met Thr Phe Arg Asn Lys Tyr Asp His Asn Tyr Ile Ala Val Cys
244 80 85 90
246 atc ttg gat tcc tct gaa aag gat tgt tat gat ttg aaa cat aac tcg 940
247 Ile Leu Asp Ser Ser Glu Lys Asp Cys Tyr Asp Leu Lys His Asn Ser
248 95 100 105
250 tta agg ata ata aac aaa gat gct ctt act tca gag tta acc tgc tta 988
251 Leu Arg Ile Ile Asn Lys Asp Ala Leu Thr Ser Glu Leu Thr Cys Leu
252 110 115 120
254 act gtt gat caa gct ttt att aac ata ccc att gaa tta ttt ggt aaa 1036
255 Thr Val Asp Gln Ala Phe Ile Asn Ile Pro Ile Glu Leu Phe Gly Lys
256 125 130 135 140
258 tac caa ata caa gat att att aat gac att gaa gca atg gga gtg att 1084
259 Tyr Gln Ile Gln Asp Ile Ile Asn Asp Ile Glu Ala Met Gly Val Ile
260 145 150 155
262 gtc aat gtt aat gta gag gca ctt agc ttt gat aat ata gga gaa aag 1132
263 Val Asn Val Asn Val Glu Ala Leu Ser Phe Asp Asn Ile Gly Glu Lys
264 160 165 170
266 cga atc caa act ttt gaa gga tat agt gtt att aca tat tct atg aaa 1180
267 Arg Ile Gln Thr Phe Glu Gly Tyr Ser Val Ile Thr Tyr Ser Met Lys
268 175 180 185
270 ttc tat aaa tat agt cac ctt ata gca aaa cga ttt ttg gat atc atg 1228
271 Phe Tyr Lys Tyr Ser His Leu Ile Ala Lys Arg Phe Leu Asp Ile Met
272 190 195 200
274 ggt gct att ata ggt ttg ctc ata tgt ggc att gtg gca att ttt cta 1276

```

RAW SEQUENCE LISTING

DATE: 10/15/2001

PATENT APPLICATION: US/09/900,038A

TIME: 13:34:08

Input Set : A:\766.53 Sequence Listing.txt

Output Set: N:\CRF3\10152001\I900038A.raw

```

275 Gly Ala Ile Ile Gly Leu Leu Ile Cys Gly Ile Val Ala Ile Phe Leu
276 205                210                215                220
278 gtt ccg caa atc aga aaa gat ggt gga ccg gct atc ttt tct caa aat 1324
279 Val Pro Gln Ile Arg Lys Asp Gly Gly Pro Ala Ile Phe Ser Gln Asn
280                225                230                235
282 aga gta ggt cgt aat ggt agg att ttt aga ttc tat aaa ttc aga tca 1372
283 Arg Val Gly Arg Asn Gly Arg Ile Phe Arg Phe Tyr Lys Phe Arg Ser
284                240                245                250
286 atg cga gta gat gca gaa caa att aag aaa gat tta tta gtt cac aat 1420
287 Met Arg Val Asp Ala Glu Gln Ile Lys Lys Asp Leu Leu Val His Asn
288                255                260                265
290 caa atg acg ggg cta atg ttt aag tta gac gat gat cct aga att act 1468
291 Gln Met Thr Gly Leu Met Phe Lys Leu Asp Asp Asp Pro Arg Ile Thr
292                270                275                280
294 aaa ata gga aaa ttt att cga aaa aca agc ata gat gag ttg cct caa 1516
295 Lys Ile Gly Lys Phe Ile Arg Lys Thr Ser Ile Asp Glu Leu Pro Gln
296 285                290                295                300
298 ttc tat aat gtt tta aaa ggt gat atg agt tta gta gga aca cgc cct 1564
299 Phe Tyr Asn Val Leu Lys Gly Asp Met Ser Leu Val Gly Thr Arg Pro
300                305                310                315
302 ccc aca gtt gat gaa tat gaa aag tat aat tca acg cag aag cga cgc 1612
303 Pro Thr Val Asp Glu Tyr Glu Lys Tyr Asn Ser Thr Gln Lys Arg Arg
304                320                325                330
306 ctt agt ttt aag cca gga atc act ggt ttg tgg caa ata tct ggt aga 1660
307 Leu Ser Phe Lys Pro Gly Ile Thr Gly Leu Trp Gln Ile Ser Gly Arg
308                335                340                345
310 aat aat att act gat ttt gat gaa atc gta aag tta gat gtt caa tat 1708
311 Asn Asn Ile Thr Asp Phe Asp Glu Ile Val Lys Leu Asp Val Gln Tyr
312                350                355                360
314 atc aat gaa tgg tct att tgg tca gat att aag att att ctc cta acg 1756
315 Ile Asn Glu Trp Ser Ile Trp Ser Asp Ile Lys Ile Ile Leu Leu Thr
316 365                370                375                380
318 cta aag gta gtt tta ctc ggg aca gga gct aag taaaggtaag gtttgaaagg 1809
319 Leu Lys Val Val Leu Leu Gly Thr Gly Ala Lys
320                385                390
322 aatata atg aaa att tgt ctg gtt ggt tca agt ggt ggt cac cta gca 1857
323      Met Lys Ile Cys Leu Val Gly Ser Ser Gly Gly His Leu Ala
324                395                400                405
326 cac ttg aac ctt ttg aaa ccc att tgg gaa aaa gaa gat agg ttt tgg 1905
327 His Leu Asn Leu Leu Lys Pro Ile Trp Glu Lys Glu Asp Arg Phe Trp
328                410                415                420
330 gta act ttt gat aaa gaa gat gct agg agt att cta aga gaa gag att 1953
331 Val Thr Phe Asp Lys Glu Asp Ala Arg Ser Ile Leu Arg Glu Glu Ile
332                425                430                435
334 gta tat cat tgc ttc ttt cca aca aac cgt aat gtc aaa aac ttg gta 2001
335 Val Tyr His Cys Phe Phe Pro Thr Asn Arg Asn Val Lys Asn Leu Val
336                440                445                450
338 aaa aat act att cta gct ttt aag gtc ctt aga aaa gaa aga cca gat 2049
339 Lys Asn Thr Ile Leu Ala Phe Lys Val Leu Arg Lys Glu Arg Pro Asp

```

VERIFICATION SUMMARY

DATE: 10/15/2001

PATENT APPLICATION: US/09/900,038A

TIME: 13:34:09

Input Set : A:\766.53 Sequence Listing.txt

Output Set: N:\CRF3\10152001\I900038A.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:25 M:283 W: Missing Blank Line separator, <400> field identifier
L:90 M:283 W: Missing Blank Line separator, <400> field identifier
L:698 M:283 W: Missing Blank Line separator, <220> field identifier
L:700 M:283 W: Missing Blank Line separator, <400> field identifier
L:709 M:283 W: Missing Blank Line separator, <220> field identifier
L:711 M:283 W: Missing Blank Line separator, <400> field identifier
L:719 M:283 W: Missing Blank Line separator, <220> field identifier
L:721 M:283 W: Missing Blank Line separator, <400> field identifier
L:729 M:283 W: Missing Blank Line separator, <220> field identifier
L:731 M:283 W: Missing Blank Line separator, <400> field identifier
L:739 M:283 W: Missing Blank Line separator, <220> field identifier
L:741 M:283 W: Missing Blank Line separator, <400> field identifier